Title
Compressed intramolecular dispersion interactions.

Journal
The Journal of chemical physics, 152(2)

ISSN
0021-9606

Authors
Mackie, Cameron J
Gonthier, Jérôme F
Head-Gordon, Martin

Publication Date
2020

DOI
10.1063/1.5126716

Peer reviewed

Compressed intramolecular dispersion interactions

Cameron J. Mackie,[1, a)] Jérôme F. Gonthier,[1] and Martin Head–Gordon[1, b)]
Kenneth S. Pitzer Center for Theoretical Chemistry, Department of Chemistry, University of California, Berkeley, California 94720, USA and Chemical Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA

(Dated: 9 March 2020)

The feasibility of the compression of localized virtual orbitals are explored in the context of *intra*molecular long–range dispersion interactions. Singular value decomposition (SVD) of coupled cluster doubles amplitudes associated with the dispersion interactions are analyzed for a number of long–chain systems, including saturated and unsaturated hydrocarbons and a silane chain. Further decomposition of the most important amplitudes obtained from these SVDs allows for the analysis of the dispersion—specific virtual orbitals that are naturally localized. Consistent with previous work on intermolecular dispersion interactions in dimers it is found that three important geminals arise and account for the majority of dispersion interactions at long range, even in the many body intramolecular case. Furthermore, it is shown that as few as three localized virtual orbitals per occupied orbital can be enough to capture all pairwise long–range dispersion interactions within a molecule.

I. INTRODUCTION

Dispersion forces dominate the long–range, non–covalent interactions between medium–to–large–sized molecules; as well as long–range intramolecular interactions within a molecule. Traditional semi–local density functional theory (DFT) methods do not describe dispersion intrinsically[1]. However substantial progress has been made by supplementing such functionals with dispersion–specific van der Waals (vdW) corrections. Prominent examples include non–local correlation functionals, such as vdW–DF–04[2] and VV10[3], and empirical atom–atom dispersion potentials such as DFT–D3[4,5], DFT–TS[6] and XDM[7,8]. The best of these functionals are very successful, but of course all such functionals remain approximate and therefore may fail in some cases, e.g., VdW B3LYP ice sinks in water[9].

In contrast, post–Hartree–Fock methods such as coupled–cluster (CC)[10], or Møller–Plesset perturbation theory[11] (MP) naturally account for dispersion interactions (vide infra), and can provide a more realistic description of dispersion than the vdW DFT methods listed above. Lowest order MP2 theory is well–known to be inadequate for relatively strong dispersion interactions such as π-stacking[12], while higher order CC methods remedy such failures.

However, these methods come at a high computational cost, and are therefore not feasible for large systems where dispersion forces can play significant roles. Combinations of DFT methods with post–Hartree–Fock methods, such as MP2 and the random phase approximation[13] (RPA), have been put forth whereby the correlation and dispersion energies can be formulated using double–excitation amplitudes calculated from previously optimized DFT orbitals[14,15]. These so–called "rung 5" functionals represent the best methods on Perdew et al.'s Jacob's Ladder to Chemical Accuracy Heaven[16,17], and come at a reasonable cost compared to full post–Hartree–Fock methods. There are promising examples that perform well for non–covalent interactions[18,19].

The application of CC methods to challenging dispersion–containing problems requires the development of more computationally efficient implementations. The most intensively pursued directions have been so–called local correlation methods[20–22], and F12 methods[23,24] that enable more rapid basis set convergence. One relatively neglected question is whether or not basis set requirements differ significantly for short–range versus long–range correlation. Recently, it has been shown that the coupled cluster double–excitation amplitudes responsible for the dispersion energy (T_2^{disp}) between dimers can be greatly compressed to as few as three virtual orbitals per occupied orbital[25]. This result fulfills the prediction made by Pulay in his first paper on local correlation ideas[26]. A virtual space compression of this type could provide a dramatic speed–up in calculation of dispersion energy calculations of dimers if incorporated into post–HF methods.

Before such efforts are attempted there are some significant open questions that should be addressed by further numerical experiments. How do the results for dimers transfer to intermolecular interactions where multiple fragments are involved? In particular, are different dispersion–specific virtuals required to capture each interaction, or are they "universal" for the interaction of a given molecule with any other species? Furthermore, since dispersion is an interaction that occurs not only between molecules, but within a molecule, how do the results for dimers transfer to interactions between functional groups within a molecule? Our goal in this work is to explore these questions by examining dispersion interactions in trimers and beyond, as well as *intra*molecular dispersion interactions between functional groups in some saturated and unsaturated hydrocarbon chain molecules.

In this work we investigate the tensor decomposition of

[a)] mackie@lbl.gov
[b)] mhg@cchem.berkeley.edu

the T_2^{disp} amplitudes through singular value decomposition (SVD) techniques in order to characterize the importance of each virtual orbital in the dispersion description. Section II briefly outlines the underlying theory of dispersion and its connection to coupled cluster amplitudes; as well as the methods used to decompose these amplitudes. Section III describes in detail the model systems used in the SVD analysis. Section IV presents the results of the decomposition of the T_2^{disp} amplitudes, as well as examples of the virtual orbitals which are important in dispersion interactions. Finally, in section V a summary of the main results of this work, and their implications, is presented.

II. THEORY

A. Two–body dispersion

Dispersion, being a non–classical phenomenon, is best understood as arising due to the fluctuations in the charge density due to the movement of electrons in a molecule. As these electrons move, their motions in a given molecule can become correlated with the motions of the electrons in another molecule, or at distant sites in the same molecule. This correlated motions lead to a lowering of the energy of the system, resulting in attraction.

Dispersion energies in the context of long–range perturbation theory[27] can be described as correlated excitations between occupied orbitals on a fragment A to virtual orbitals on the same fragment A; with occupied orbitals on another fragment B to the virtual orbitals on that same fragment B. Within the Hartree–Fock wavefunction framework, the long–range dispersion energy can be expressed to lowest order as

$$E_{disp} = -4 \sum_{ia \in A} \sum_{jb \in B} \frac{(ia|jb)\,(ai|bj)}{\epsilon_a + \epsilon_b - \epsilon_i - \epsilon_j} \tag{1}$$

where, $(ia|jb)$ is the molecular repulsion integral over spatial–orbitals using chemist's notation, and ϵ_x is the Hartree–Fock orbital energy of orbital x (a, b virtual; i, j occupied). Equation 1 is only valid when exchange interactions *between* fragments are negligible, a reasonable assumption to be made at medium to long range due to the exponential decay of the exchange term. Within this assumption, i and a become the occupied and virtual orbitals of fragment A respectively, and correspondingly j and b become the occupied and virtual orbitals of fragment B.

Equation 1 can be rewritten as

$$E_{MP2}^{2-body\ disp.} = 4 \sum_{ia \in A} \sum_{jb \in B} t_{ij}^{ab}\,(ai|bj) \tag{2}$$

where t_{ij}^{ab} are doubles amplitudes. The MP2 doubles at long range are:

$$t_{ij}^{ab} = -\frac{(ia|jb)}{\epsilon_a + \epsilon_b - \epsilon_i - \epsilon_j} \tag{3}$$

(subject to the same restrictions given on fragments A and B above). Since MP2 amplitudes are known to over estimate dispersion energies in the basis set limit[28,29], the t_{ij}^{ab} amplitudes in equation 2 can be substituted with the more accurate CCSD amplitudes, similar to previous studies on the energy decomposition of the full exchange–dispersion terms[30–33] and the previous SVD analysis of dimers[25] (see ref. 25 for more details).

B. Three–body dispersion

At the MP2 level, there is no three–body dispersion, because equation 1 is pairwise additive. At the MP3 level *disconnected*–triples are included, i.e.,

$$\begin{aligned} E_{MP3}^{3body\ disp.} = 16 \sum_{ia \in A} \sum_{jb \in B} \sum_{kc \in C} \\ t_{ij}^{ab} t_{ik}^{ac}\,(bj|ck) \\ + t_{ji}^{ba} t_{jk}^{bc}\,(ai|ck) \\ + t_{ki}^{ca} t_{kj}^{cb}\,(ai|bj) \end{aligned} \tag{4}$$

Likewise, at the CCSD level disconnected triples also make an appearance through the relaxation of the doubles amplitudes in the presence of the third body, which then indirectly affects the total CCSD energy via equation 2.

Analogous to equation 1 the *connected* triples and their associated three body dispersion energy can be written as

$$\begin{aligned} E_{MP4}^{3body\ disp.} = -\frac{4}{3} \sum_{ia \in A} \sum_{jb \in B} \sum_{kc \in C} \\ \frac{|w_{ijk}^{abc}|^2}{\epsilon_a + \epsilon_b + \epsilon_c - \epsilon_i - \epsilon_j - \epsilon_k} \end{aligned} \tag{5}$$

where the triples–substitution term w_{ijk}^{abc} is defined in ref. 34. w_{ijk}^{abc} can be expressed in the fragment based spatial–orbital form used in this work as

$$
\begin{aligned}
w_{ijk}^{abc} = & \sum_{e}^{virt\in A} t_{ij}^{eb}\,(ae|ck) + t_{ik}^{ec}\,(ae|bj) \\
& + \sum_{f}^{virt\in B} t_{ij}^{af}\,(bf|ck) + t_{jk}^{fc}\,(bf|ai) \\
& + \sum_{g}^{virt\in C} t_{jk}^{bg}\,(cg|ai) + t_{ik}^{ag}\,(cg|bj) \\
& - \left(\sum_{m}^{occ\in A} t_{mj}^{ab}\,(mi|ck) + t_{mk}^{ac}\,(mi|bj) \right) \\
& - \left(\sum_{n}^{occ\in B} t_{in}^{ab}\,(nj|ck) + t_{nk}^{bc}\,(nj|ai)) \right) \\
& - \left(\sum_{p}^{occ\in C} t_{jp}^{bc}\,(pk|ai) + t_{ip}^{ac}\,(pk|bj)) \right)
\end{aligned}
\tag{6}
$$

Note that it is necessary to increase to fourth order perturbation theory to account for *connected*–triples. Similarly in coupled–cluster, triple excitations (i.e, CCSDT) are needed to account for the connected–triples terms of the three–body dispersion energy. See references 35–37 for descriptions of many–body interactions, and reference 38 and references therein for a full description and benchmarks of three–body dispersion in small clusters.

Fortunately, three–body dispersion decays as R^{-9} so its contribution at long range is small. (Although, one should be cautious of this assumption in large systems such as bulk liquids or solids where these three–body dispersion interactions add up quickly[39–43].) However, in this work, the tiny error introduced in the neglect of connected–triples would be quickly overwhelmed by any error introduced in the compression of the virtual space itself. For this reason, as well as computational tractability, we will limit ourselves to the CCSD doubles amplitudes in this work. Therefore, the many–body contributions to dispersion will enter via relaxation of the CCSD amplitudes. As such, this work will concern pairwise additive three–body dispersion, and neglect non–additive three–body dispersion.

C. Two–fragment Singular Value Decomposition

The analysis of the compressibility of dispersion interactions between two monomers follows the method described in our previous work[25]. First, we choose a suitable tensor describing dispersion interactions, for example the CCSD $\mathcal{T}_2$ amplitudes. To select the intermolecular block of the tensor, and thus separate the analysis of intermolecular correlation effects from intramolecular correlation effects, we localize orbitals using the Boys localization procedure. At sufficiently large intermolecular separation (found to be sufficient beyond 3 Å in reference 25 and 2.6 Å in this work), the orthogonalization tails that result from this process are negligible. The resulting dispersion tensor to analyze is denoted $\mathcal{T}_2^{disp}$.

Singular Value Decomposition (SVD) yields the best rank–reduced approximation of a matrix in a least–square sense. Therefore, the decomposition of the four–index tensor $\mathcal{T}_2^{disp}$ is performed by first flattening the tensor to two dimensions. In our case, we are specifically interested in extracting monomer–related properties, thus we grouped together indices from the same monomer: $\mathcal{T}_2^{disp}$ elements are arranged in a matrix $\boldsymbol{T}^{O_A V_A, O_B V_B}$ where the occupied and virtual indices of one monomer run along the rows while those of the other monomer run along the columns. Applying SVD to this matrix yields:

$$
\boldsymbol{T}^{O_A V_A, O_B V_B} = \boldsymbol{G}^A \boldsymbol{\Gamma} \left(\boldsymbol{G}^B\right)^T \tag{7}
$$

where we associate the left singular vectors $\boldsymbol{G}^A$ with monomer A and the right singular vectors $\boldsymbol{G}^B$ with monomer B. The diagonal matrix $\boldsymbol{\Gamma}$ contains the singular values, which rank the importance of each pair of singular vectors in the original matrix. Physically, we interpret the SVD as expressing a tensor of double excitations into a sum of products of single excitations, each of which occurs on a single monomer. Note that each single excitation involves only one electron but may contain components from different orbitals, as we will see below. We can thus rewrite the above equation to explicit each single excitation as a column vector $\boldsymbol{G}_{\bullet P}$, with γ_P the weight of each pair of excitation:

$$
\boldsymbol{T}^{O_A V_A, O_B V_B} = \sum_{P}^{N_{gem}} \boldsymbol{G}_{\bullet P}^A \cdot \gamma_P \cdot \left(\boldsymbol{G}_{\bullet P}^B\right)^T \tag{8}
$$

This interpretation is consistent with the dimensions of the singular vectors. Since each of those vectors contain information about both the occupied and the virtual space of each monomer, we refer to them as geminal vectors. This first SVD already yields the effective rank of the $\boldsymbol{T}^{O_A V_A, O_B V_B}$ matrix since we can eliminate pairs of geminals associated with negligible weights γ_P.

To gain further insight into the significant excitation processes in dispersion interactions, we extract information from each geminal vector $\boldsymbol{G}_{\bullet P}$. In particular, we can separate the occupied from the virtual space in each excitation by applying SVD again, in a similar spirit to the Natural Transition Orbital[44] analysis. Each vector is reshaped as a matrix $\left(\boldsymbol{G}_{\bullet P}^X\right)^{O_X, V_X}$ where occupied and virtual indices are arranged along the rows and columns, respectively. Singular value decomposition then yields:

$$
\left(\boldsymbol{G}_{\bullet P}^X\right)^{O_X, V_X} = \boldsymbol{U} \boldsymbol{\Gamma} \left(\boldsymbol{V}\right)^T \tag{9}
$$

where the single excitation associated with the geminal vector is now expressed as a sum of occupied to virtual pairs, weighted by the singular values in Γ. The matrices $\boldsymbol{U}$ and $\boldsymbol{V}$ rotate the occupied and the virtual space respectively, yielding orbitals that most efficiently describe

the original single excitation. Since our interest usually lies in truncations of the virtual space, we restrict our analysis to matrix $\boldsymbol{V}$. By transforming the significant columns of this matrix to the Atomic Orbital basis, we obtain and plot the most important virtual orbitals involved in the excitation processes giving rise to dispersion interactions[25].

D. Multi–fragment SVD

For the multi–fragment SVD analysis the T_2^{disp} tensor is unfolded into a matrix similar to the two fragment SVD analysis described above. However, in order to account for the multiple simultaneous interactions, the tensors are unfolded into *interaction–blocks*. Thus for a system containing fragments A, B, C, etc, we may consider the $\mathbf{T}_2^{disp}$ tensor matrix to be composed as follows:

$$\mathbf{T}_2^{disp} = \begin{bmatrix} \mathbf{0} & \mathbf{T}^{AB} & \mathbf{T}^{AC} & \cdots \\ \mathbf{0} & \mathbf{0} & \mathbf{T}^{BC} & \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \\ \vdots & & & \ddots \end{bmatrix} \tag{10}$$

Here we understand that $\mathbf{T}^{AB}$ represents $\mathbf{T}^{\mathbf{O_A V_A}, \mathbf{O_B V_B}}$, etc. The diagonal blocks are set to zero, because there is no self–dispersion, and lower–triangle blocks are also set to zero because they are duplicates of the non–redundant upper triangular blocks.

III. MODEL SYSTEMS

Multiple systems were analyzed to test the compressibility of the T_2^{disp} amplitudes. All molecular geometries were optimized using DFT[45,46] at the ωB97X–D[47]/6–31G*[48] level of theory. The T_2^{disp} amplitudes were obtained at these geometries using the CCSD[49] method with the cc–pVDZ basis set[50], with core electrons frozen. Due to the number of electrons in the largest molecules analyzed, cc–pVDZ was the largest basis set that could be used. As a cross–check, n–hexane was also run with the aug–cc–pVDZ and cc–pVTZ[50] basis–set to assess basis–set dependence. All calculations were performed using a locally modified version of the Q–Chem software package[51].

A. Helium trimer

The helium trimer was chosen for a full multi–fragment SVD analysis to first establish the number of important geminals for three–body pair–wise additive dispersion interactions. As will be shown below, the SVD of helium is also used to justify the splitting of the full multi–fragment unfolded T_2^{disp} tensor into its sub–block components for individual SVD analysis. Two geometries are utilized, as shown in figure 1: an equilateral triangular geometry, where the inter–atomic distances are increased equally, and a linear geometry, where one inter–atomic distance is fixed at 3 Å and the other is variable.

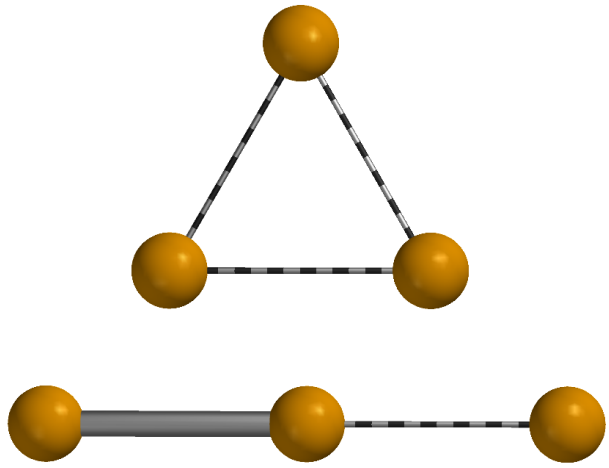

FIG. 1: Two helium trimer configurations. The checkered lines represent inter–atomic distances which are increased during the SVD analysis, while the solid line represents a fixed inter–atomic distance of 3 Å

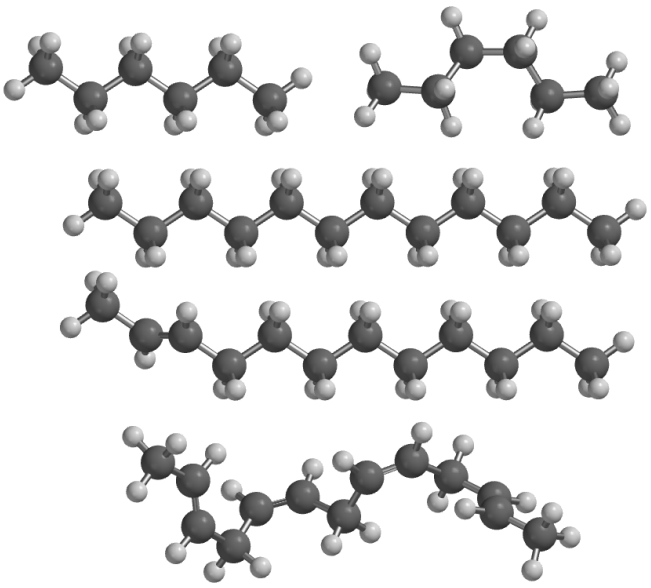

FIG. 2: Model hydrocarbon systems used in the intramolecular analysis, n–hexane, gauche hexane, n–dodecane, trans–2–dodecene, 2,5,8,11–tridecene. Carbon bonds and CH–groups are enumerated from left to right as shown during the SVD analysis.

B. Hydrocarbons and silanes

After establishing the validity of the pair–wise sub–block SVD analysis (vide infra) it became possible to investigate intramolecular dispersion. A series of hydrocarbon molecules were chosen to test the compressibility of the intramolecular T_2^{disp} amplitudes: n–hexane, gauche hexane, n–dodecane, trans–2–dodecene, and 2,5,8,11–tridecene (see figure 2). The occupied and virtual orbitals were localized separately using the Boys algorithm[52–55]. This resulted in two active electrons localized in each bond–centered occupied orbital, as well as a series of bond–centered virtual orbitals for each bond.

n–Hexane

n–Hexane was chosen as it is the simplest hydrocarbon to test long–range intramolecular dispersion. A question arises as to what to consider a "fragment" within a molecule for the SVD analysis. To address this, n–Hexane was analyzed with both CC–bonds constituting a fragments as well as CH–groups constituting a fragment (all other hydrocarbons are analyzed with CC–bonds only). The SVD analysis of n–hexane was also performed with the aug–cc–pVDZ and cc–pVTZ basis–set amplitudes to explore basis–set dependence. The CC–bond numbering given in figure 5 refers to every other CC–bond, starting from the left as shown in figure 2 (hydrogen atoms are purely "spectators"). Likewise, the CH–group numbering refers to each CH–group starting on the left with CH_3, then CH_2, etc.

Gauche hexane

A gauche configuration of hexane was chosen to compare with the results of n–hexane to test for possible differences between "through–bond" dispersion versus "through–space" dispersion, and to break any possible contributions from hyperconjugation. CC–bond numbering in figure 5 remains the same as n–hexane.

n–Dodecane

n–Dodecane was chosen in order to measure the distance dependence of the singular values. As distance increases the dispersion energy for dipole–dipole interactions are expected to decay as R^{-6}; for dipole–quadrupole interactions decays as R^{-8}; and for quadrupole–quadrupole interactions decays as R^{-10}, which correspond to singular value decays of R^{-3}, R^{-4}, and R^{-5} respectively. The resulting energies at various degrees of compression of dispersion–specific orbitals for the CC–bond interactions are also analyzed. CC–bond numbering in figure 5 refers to every other CC–bond as shown in figure 2, starting with the second CC–bond (to facilitate easier comparison with trans–2–dodecene below).

trans–2–Dodecene

trans–2–Dodecene was chosen in order to determine if the SVD compressibility results hold for π–bonded systems. Additionally, the distance dependence of the singular values between a double–bond and single–bonds is examined. CC–bond numbering in figure 5 refers to every other CC–bond as shown in figure 2, starting with the second CC–bond. The second CC–bond being a double–bond, with all other bonds being single–bonds.

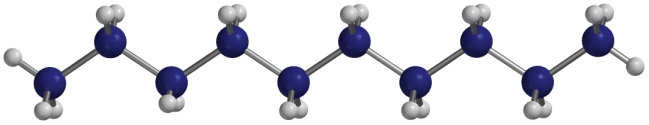

FIG. 3: Model silane system used in the intramolecular analysis, Si_{10}–silane.

2,5,8,11–Tridecene

2,5,8,11–Tridecene was chosen in order to determine the effect which π–π interactions have on the SVD compressibility, as well as determine the relative importance of π versus σ dispersion. Each double–bond is separated by two single–bonds in order to avoid conjugation effects, with the first double–bond occurring between carbon two and three. A low–symmetry, random orientation of the dihedral was employed to increase the variance of singular values. An unintended consequence of this random orientation is that virtual orbital orientation can be examined with regards to dispersion interactions. The CC–bond numbering used in figure 5 refers to each double–bond in order from left to right.

Long–chain silane

Si_{10}–silane was chosen as an analog to the hydrocarbons in this work in order to test the transferability of these SVD findings to non–hydrocarbon systems. As in n–hexane, both the SiSi–bonds and SiH–groups were analyzed. The SiSi–bond numbering given in figure 6 refers to every other SiSi–bond, starting from the left as shown in figure 3. Likewise, the SiH_x–group numbering refers to each SiH_x–group starting on the left with SiH_3, then SiH_2, etc.

IV. RESULTS

A. Singular value decomposition

Helium trimer

The full multi–fragment SVD results for the helium trimer at three atomic distances is presented in figure 4. The magnitudes of the largest singular values are shown. The top panel shows the singular values for the equilateral triangular configuration, and the bottom for the linear configuration (see figure 1). As can be seen for the triangular case, there are nine important geminals of roughly equal magnitude before a sharp drop–off of over two orders of magnitude for the 3 Å distance, and well over four orders of magnitude for 6 and 9 Å. As shown in reference 25, this sharp drop–off leads to a high compressibility of the $\mathbf{T}_2^{disp}$ terms used to calculate the dispersion energy with little loss to accuracy. Equally important in the context of this work, is that exactly nine important geminals are found, i.e., three geminals per significant pairwise interaction. This nicely generalizes the dimer SVD analysis of reference 25, where 3

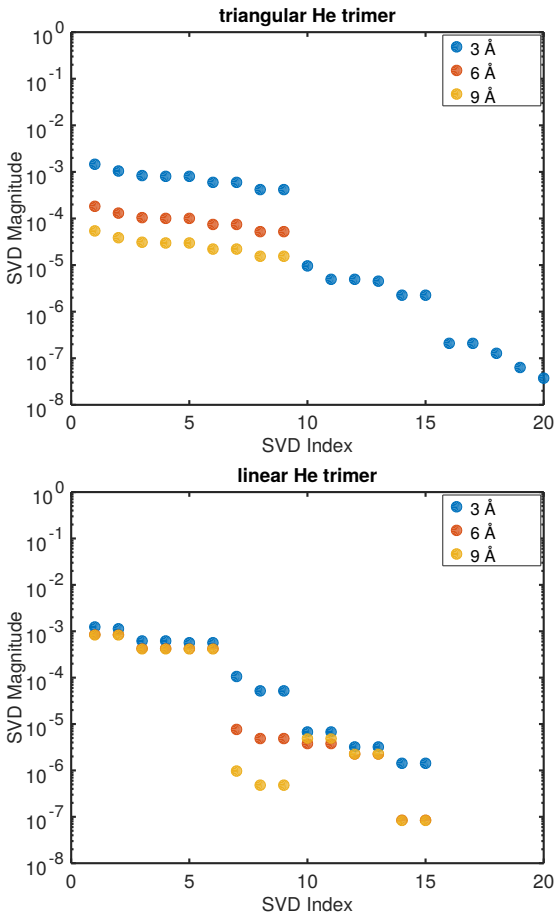

FIG. 4: The $\mathbf{T}_2^{disp}$ singular values for the He trimer interactions. The upper panel corresponds to SVD at the equilateral triangle geometry, and the lower panel to SVD at the linear geometry: see figure 1 for the structures and bond definitions. The ordering of the singular values are altered in the 9Å case of second panel so as to reflect the decay progression of the relevant singular values.

.

significant geminals were found for a single interaction in a helium dimer. Confirmation of this can be seen in the linear He trimer case shown in the lower panel of figure 4. As one He atom is pulled further away from the other two, only three geminals are shown to decrease appreciably, corresponding to a weakening interaction.

The result that the $\mathbf{T}_2^{disp}$ tensor for N interacting monomers is represented to leading order by only $3N(N-1)$ geminals simplifies the examination of more complex systems. No specific three–body virtual orbitals need to be considered. The lack of any significant contribution from geminals representing three–body terms in the decomposition also leads to the conclusion that the SVD analysis can be carried out simply as pairwise, i.e., the sub–blocks of the unfolded $\mathbf{T}_2^{disp}$ can be decomposed and analyzed independently from one another. This opens up the possibility to examine the dispersion contributions of the amplitudes in large systems in an efficient manner, including intramolecular dispersion interactions.

Furthermore, the second SVD of the $3N(N-1)$ important geminals into occupied–virtual pairs reveals redundancy in the virtual space. That is to say, the same occupied–virtual pairs per monomer are responsible for the dominant dispersion interactions for all other fragments. This further reduces the number of important dispersion–specific virtual orbitals per occupied from as many as $3N(N-1)$ to as few as $3N$. A more thorough analysis of this reduction in virtual space is left to the more complex hydrocarbon systems in section IV E.

Hydrocarbons and silane

As stated in section III A the SVD of the hydrocarbon intramolecular dispersion is performed on the sub–blocks of the unfolded $\mathbf{T}_2^{disp}$ tensor. Descriptions of the CC–bond and CH–group number of figure 5 are given in section III B. Only the SVD between the first bond/group and all other bonds/groups are shown in each plot, similar results were obtained for all other intramolecular interactions.

Hexane

Hexane represents the smallest hydrocarbon in which the compressibility of the intramolecular dispersion $\mathbf{T}_2^{disp}$ amplitudes becomes apparent. The top–left panel of figure 5 shows the singular value magnitudes for the CC–bond:CC–bond dispersion of n–hexane. For CC–bond 1 on CC–bond 3, the drop–off in singular value between the third and fourth singular value is marginal. This is expected since the distance between these bonds is approximately 2.6 Å, well below the distance where long–range dispersion would dominate, and likely in a region where exchange interactions are significant (violating the conditions for equation II A). Moving further out to the SVD of CC–bond 1 with CC–bond 5 (a distance of approximately 5.1 Å) the drop–off between the third and fourth singular value becomes more pronounced, representing an order of magnitude decrease. Interestingly, a drop–off between the first and second singular value also appears, again about an order of magnitude, which will be discussed in more detail in the context of distance dependence below.

For n–hexane an SVD analysis which makes use of the CH–groups for the fragments is also presented in the top panel of figure 5. Again, a drop–off between singular values three and four is observed, which increases with fragment distance. The drop–off reaches an order of magnitude between CH–group 1 and CH–group 4, representing an inter–fragment distance of approximately 5.1 Å, consistent with the CC–bond n–hexane SVD.

As stated in section III B a gauche hexane configuration was also explored to test for differences between through–bond and through–space dispersion, and to break any possible contributions from hyperconjugation. As can be seen in the middle left panel of figure 5 compared to n–hexane in the top left, little changes. The drop–offs are slightly diminished, and the gap between

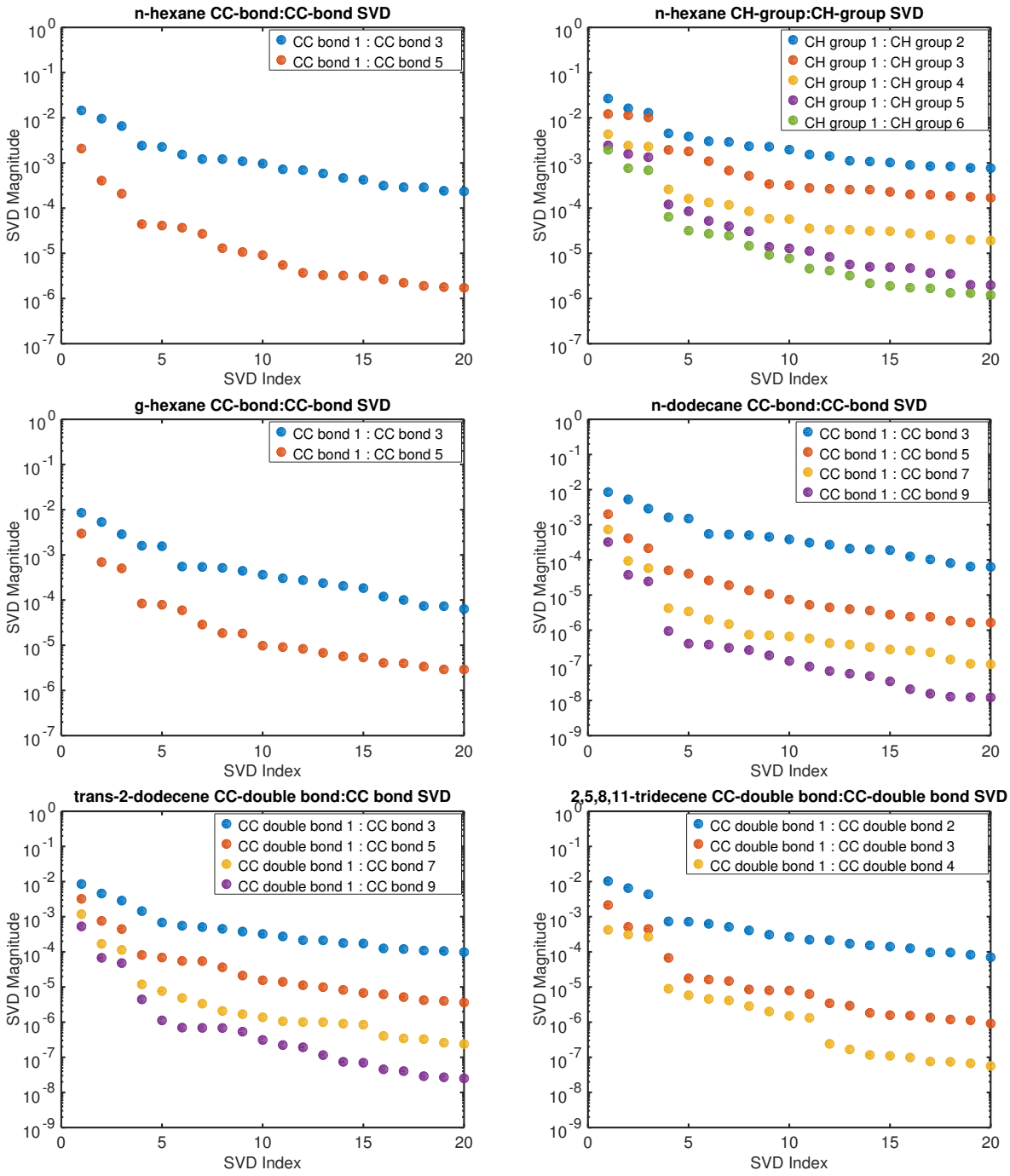

FIG. 5: The first twenty $\mathbf{T}_2^{disp}$ singular values for the bond–bond dispersion interactions of four long–chain hydrocarbons: n–hexane (CC–bonds), n-hexane (CH–groups), g–hexane, n–dodecane, trans–2–dodecene, and 2,5,8,11–tridecene. See figure 2 for the relevant structures and bond definitions.

CC–bond 1 on CC–bond 3 and CC–bond 1 on CC–bond 5 is closed, however this is due likely to the shortening of the CC–bond CC–bond distances when moving from the normal to gauche configuration, although orbital orientation does play a role (albeit minor in this case) as will be shown with 2,5,8,11–tridecene below.

n–Dodecane

Qualitatively, the singular values of n–dodecane shown in the middle–right panel of figure 5 follow the pattern seen in n–hexane. At short distances i.e., between CC–bond 1 and CC–bond 3 (2.6 Å), the drop–off between the third and fourth singular value is not apparent. Between CC–bond 1 and CC–bond 5 (5.1 Å) the drop–off appears, then continues to grow between CC–bond 1 and CC–bond 7

(7.7 Å), and CC–bond 1 and CC–bond 9 (10.2 Å).

trans–2–Dodecene

As for n–dodecane, trans–2–dodecene was chosen to examine the distance dependence of the singular values; in this case to examine specifically the decay rate of the dispersion interaction between a CC–double–bond and CC–single–bonds. Again, qualitatively for the first three singular values trans–2–dodecene is consistent with n–hexane and n–dodecane in that they show a clear drop–off after the third singular value which expands with distance, and the first singular value dominates. However, the drop–off between the third and fourth singular value is not as large in magnitude (clearly visible between CC–double–bond 1 and CC–double–bond 4 (10.2 Å)), before again dropping off at the fifth singular value.

2,5,8,11–Tridecene

2,5,8,11–Tridecene was originally selected in order to examine π/σ–bond on π/σ–bond dispersion interactions and the relative importance of π versus σ contributions. However, the bottom–right panel of figure 5 shows clear erratic behavior when moving out to further bonds. Upon inspection of the virtual orbital components of the most important geminals it becomes clear that orientation of the orbitals play a strong role (see section IV E).

Si_{10}–silane

Si_{10}–silane was selected as a non–hydrocarbon analog to the long–chain hydrocarbons in this work. The behavior of the singular values remain largely the same. The top panel of figure 6 shows the singular values of the $\mathcal{T}_2^{disp}$ of the Si_{10}–silane bond–bond dispersion. The first three singular values remain the most important at long–distances, with an enhancement of the first singular value, however the drop–off does not become clear until Si–bond 1 on Si–bond 7 (11.9 Å). This delayed drop–off is likely due to the larger exchange correlation effects of the extended orbitals of the Si atoms. It is also possible that much like in the case of trans–2–dodecene, the fourth singular value sees an enhancement. As such, if the fourth singular value is ignored, the drop–off becomes significant at Si–bond 1 on Si–bond 5 (7.9 Å). Although appealing (especially if one examines the fourth singular value of Si–bond 1 on Si–bond 7) the evidence is to the contrary. Table I shows no R^{-4} dependence for the fourth singular value, and inspection of the virtual orbitals shows no dipole–quadrupole interactions.

As with n–hexane the SVD of Si_{10}–silane was also carried out using the SiH–groups, as shown in the bottom panel of figure 6. The results are consistent with n–hexane, a steep drop–off after the third singular value. No odd behaviors as seen in the Si–bond SVD are observed. The drop–off become significant much sooner than the Si–bond SVD as well; occurring at SiH group 1 on SiH group 3 (4.0 Å).

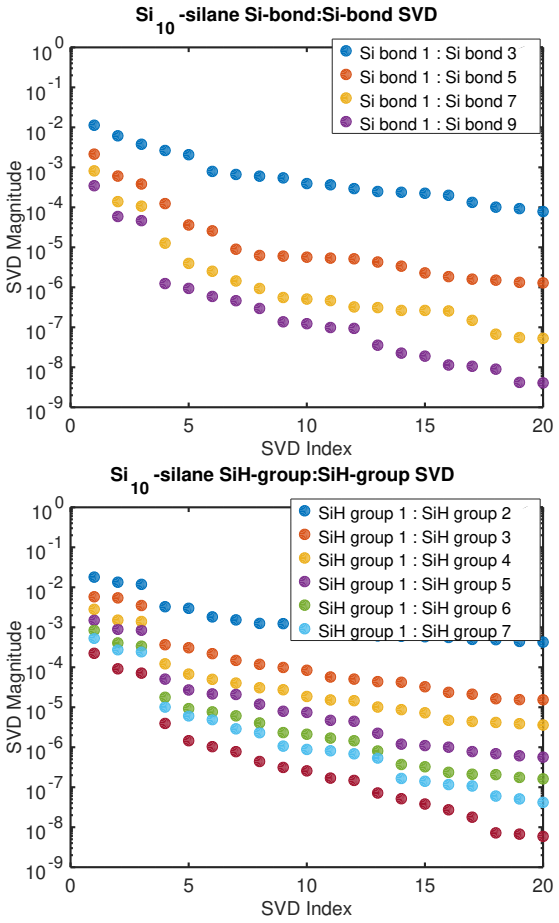

FIG. 6: The $\mathcal{T}_2^{disp}$ singular values for the bond–bond dispersion interactions of the long–chain silane. Si–Si bond interactions are shown in the top panel, while SiH–SiH group interactions are shown in the bottom panel. See figure 2 for the relevant structure and bond definitions.

B. Basis–set dependence

n–Hexane is small enough to allow comparison with larger basis sets. An SVD analysis was performed using the cc–pVDZ, aug–cc-pVDZ, and cc–pVTZ basis sets. Figure 7 shows the singular values of both the CC–bond 1 on CC–bond 3 (top panel) and CC–bond 1 on CC–bond 5 (bottom panel) dispersion interactions for each basis. Small differences between the three basis–sets are observed, ruling out any strong basis–set dependence. The sharp drop off between the third and fourth geminals are observed for all basis–sets. The differences in decay rates of the singular values occurring for the C–bond 1 on CC–bond 5 dispersion interaction are likely due to a better description of long–range diffuse virtual–orbitals. However, their contribution is negligible when compared to the contribution of the first three geminals taken together. This result points to a fast basis–set convergence at least for the construction of the first three important

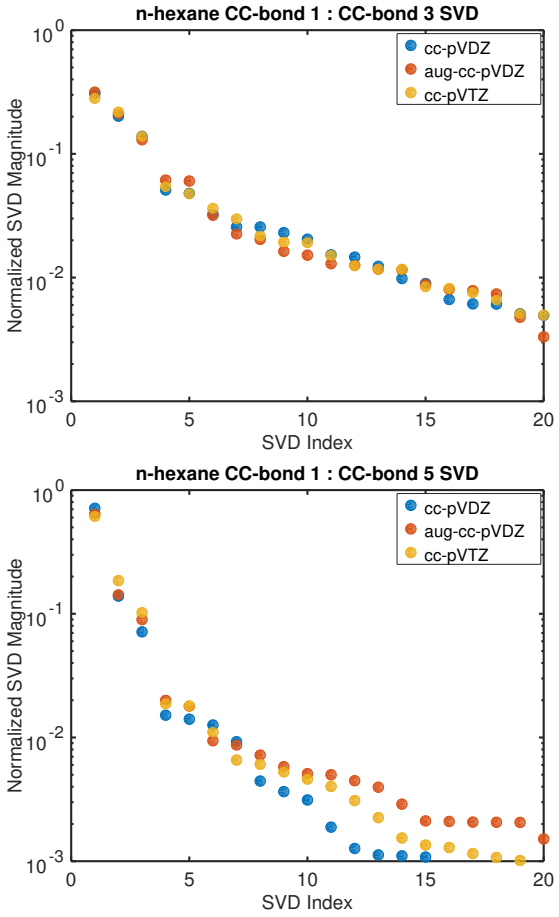

FIG. 7: The $\mathcal{T}_2^{disp}$ singular values for the CC–bond 1 on CC–bond 3 (top panel), and CC–bond 1 on CC–bond 5 (bottom panel) dispersion interactions of hexane with three basis–sets: cc–pVDZ, aug–cc–pVDZ, and cc–pVTZ.

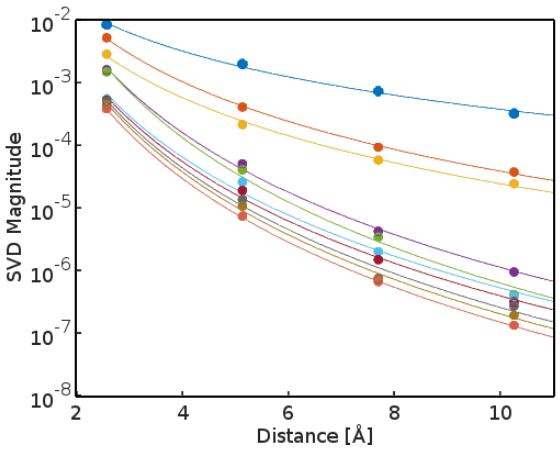

FIG. 8: Singular value decay as a function of distance in n–dodecane. Power–law fits are provided for each singular value (see table I for the corresponding exponents).

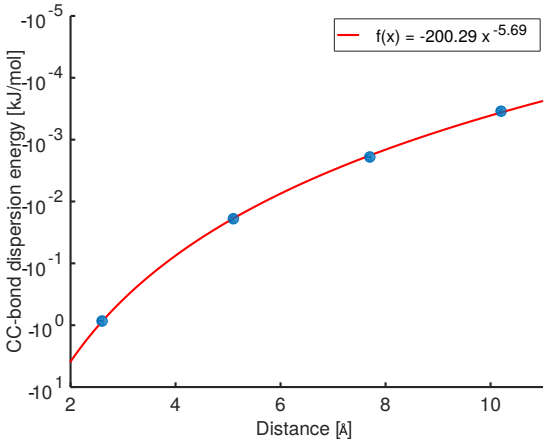

FIG. 9: The distance dependence of CC–bond – CC–bond dispersion interactions of n-dodecane. A power–law fit is given in red.

dispersion–specific virtual orbitals.

C. Distance dependence

n–Dodecane

The SVD of the $\boldsymbol{T}_2^{disp}$ amplitudes of n–dodecane are performed primarily in order to test for distance dependence of the singular values, and hence the distance dependence of the orbital contributions to the dispersion energy. The correlation energy can be computed as the product of the CCSD amplitudes (decomposed as singular values and vectors) with the two-electron integrals. To the first order, the decay behavior of the singular values and the corresponding integrals will be the same[25]. Thus, the decay of the correlation energy will be the square of the singular values' decay, i.e., if singular values decay as R^{-3}, then the energy should decay as R^{-6}.

Figure 8 shows the decay of the first ten singular values of n–dodecane with distance, with their corresponding power–law fits., and table I lists the exponents used for the fits. The first three singular values decay as approximately R^{-3} (R^{-6} in energy), consistent with the dipole–dipole contributions to dispersion. However, the first singular value is enhanced in both magnitude and slowed in decay (closer to R^{-2}) compared to the next two. As will be discussed in section IV E this enhanced interaction is due to excitations to p–like virtual orbitals whose node cuts the bond perpendicularly. The next seven singular values decay as approximately R^{-5} (R^{-10} in energy) consistent with the quadrupole–quadrupole contributions to dispersion. Interestingly, the dipole–quadrupole contribution to dispersion (R^{-4}, R^{-8} in energy) is not a strong contributor as is seen for inter–molecular dimers in reference[25].

TABLE I: Exponents for the fitting curves of the first 10 geminals of n–dodecane, trans–2–dodecene, and Si_{10}–silane; given by $f(x) = b \times x^A$ where x is a given intramolecular C–C C–C (Si) bond distance.

	SVD index	1	2	3	4	5	6	7	8	9	10
n–dodecane	A (-)	2.34	3.59	3.45	5.41	5.84	5.25	5.39	5.60	5.70	5.77
trans–2–dodecene	A (-)	2.58	3.49	3.21	4.21	5.88	6.25	6.31	5.80	5.33	5.63
Si_{10}–silane	A (-)	2.48	3.37	3.18	5.39	5.57	5.22	5.28	5.54	6.03	5.87

The distance dependence for the complete dispersion interaction (i.e., all orbital contributions) between one CC–bond and the remaining CC–bonds is given in figure 9 as a function of distance. Interestingly, the total dispersion interaction is found decay as $R^{-5.69}$, rather than the expected R^{-6}, likely due to the slower decay of the enhanced perpendicular p–like virtual–orbital mentioned above.

trans–2–Dodecene

Table I also list the exponents used for the power–law fits of the decay of the first ten singular values of trans–2–dodecene with distance. Again, the decay of the first three singular values decay as approximately R^{-3} (R^{-6} in energy), with the first singular value decaying closer (but not as close as n–dodecane) to R^{-2}. The fourth singular value decays as R^{-4} (R^{-8} in energy) consistent with a dipole–quadrupole contribution to dispersion. This explains the delay in drop–off between the third and fourth singular values seen in trans–2–dodecene. As will be shown in section IV E this is due directly to the presence of the π–orbital and its corresponding virtual orbitals. Singular values five through ten (excluding nine) also show deviating behavior. Their decays are closer to R^{-6}, with singular values six and seven even exceeding R^{-6}. Inspection of their virtual orbital components still classifies them as quadrupole–quadrupole dispersion terms. Instead *orientation* of the interacting bonds plays a role (trans–2–dodecene is not longer a straight hydrocarbon–chain). This effect is more pronounced in 2,5,8,11–tridecene and so will be left for further discussion in section IV E.

2,5,8,11–Tridecene

Although no fit to decay rates of the singular values can be made for 2,5,8,11–tridecene due to the sporadic behaviour of the singular values (see figure 5), qualitative descriptions and predictions of the singular values can still be made. The drop–off between the third and fourth singular values remains, suggesting strong dipole–dipole contributions to dispersion. However, the first singular value is enhanced only for CC–double–bond 1 on CC–double bond 3, suggesting orientation considerations. For CC–double–bond 1 on CC–double bond 3 (but not CC–double bond 2 or 4) the fourth singular value is enhanced, suggesting a contribution from a dipole–quadrupole interactions.

Si_{10}–silane

Lastly, table I shows the exponents used in the power–law fits of the decay of the first ten singular values the Si–bond interactions of Si_{10}–silane with distance. As in the hydrocarbons, the first three singular values decay as approximately R^{-3}, with the first singular value decaying slower. The next five singular values decay as approximately R^{-5} consistent with quadrupole–quadrupole contributions to dispersion. However, the last two singular values decay as approximately R^{-6}. Unlike trans–2–dodecene this change in decay rate does not reflect orbital orientation. Examination of the virtual orbitals involved in these geminals display complex forms, possibly hinting at quadrupole–hexapole interactions.

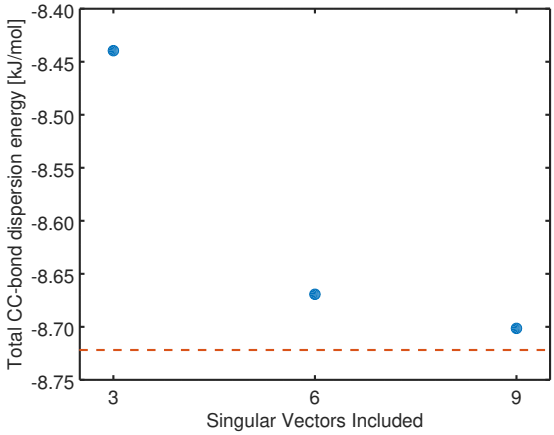

FIG. 10: Total dispersion energy between all CC–bonds contained within n–dodecane at various degrees of compression. The full uncompressed theoretical dispersion energy is given by the dashed red line.

D. Compressed intramolecular dispersion energies

Equation 2 (using CCSD amplitudes) along with the singular values obtained from equation 7, allows for the calculation of the dispersion energies at various levels of compression, i.e., various numbers of singular vectors retained (number of dispersion–specific virtual orbitals per occupied). Figure 10 shows the total intramolecular dispersion energy from all CC–bonds contained within n–dodecane at various degrees of compression. The uncompressed dispersion energy is given by the dashed red line (-8.72 kJ/mol). As can be seen, a compression down to three singular vectors, i.e., three virtual orbitals per occupied, gives a total dispersion energy of -8.44 kJ/mol, a difference of 0.28 kJ/mol, capturing 96.8% of the total

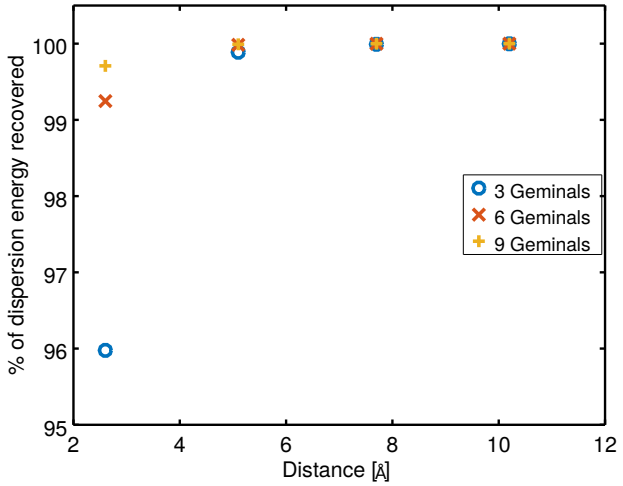

FIG. 11: The percent of the total CC–bond dispersion energy of n–dodecane recovered at various distances. Three levels of compression of the virtual space are shown, 3 virtuals per occupied (blue circles), 6 virtuals per occupied (red x's), and 9 virtuals per occupied (yellow crosses).

4.95014	4.91730	4.91419	0.08487	0.07638
0.05106	0.00402	0.00152	0.00027	0.00020
0.00005	0.00000	0.00000	0.00000	0.00000

TABLE II: Eigenvalues of the overlap matrix of n–dodecane showing the linear dependence between the virtual orbitals used to calculate dispersion interactions.

dispersion energy originating from the CC–bonds. Figure 13 provides a further breakdown of the percentage of the dispersion captured at the three levels of compression as a function of distance. Beyond 5.1 Å nearly 100% of dispersion is captured by the minimal three virtual orbitals per occupied (geminals).

E. Virtual orbitals

1. Shape of dispersion virtual orbitals

As stated in section II C the geminals obtained during the SVD analysis can be further decomposed into their occupied–virtual pairs. Examination of the virtual orbital components of each species shows a striking pattern. The largest contributing geminals correspond to virtual orbitals consisting of one higher angular momentum than the occupied orbitals they are excited from. For example, figure 12 shows the occupied and virtual components of the first five significant geminals for the π–orbital and first three significant geminals for the σ–orbital of the double–bonded carbon atoms of trans–2–dodecene. It can be seen that the π–orbital (p–like) is excited to a series of d–like orbitals, and likewise the σ–orbital (s–like) is excited to a series of p–like orbitals.

2. Orientation dependence

As alluded to previously, the orientation of the occupied and virtual orbitals affects the magnitude of dispersion interaction between two fragments. Due to the Boys bond centered localization of the orbitals employed in this work, the occupied and virtual orbitals have their angles and nodes fixed to be parallel or perpendicular to the bonds. In a non–linear or non–planar molecule bonds can be oriented relative to one another at odd angles or odd dihedral angles, and as such the virtual orbitals controlling dispersion interactions can also be at odd angles to one another resulting in stronger or weaker dispersion interactions. 2,5,8,11–Tridecene can be used to illustrate this effect. Figures **??**, 14, and 15 show examples of pairs of virtual orbitals contributing to dispersion.

The dominant geminals, $\boldsymbol{G}^A$ and $\boldsymbol{G}^B$ associated with the dispersion interaction of bonds A and B are charge distributions with zero total charge. Their leading long–range electrostatic interaction is therefore dipole–dipole:

$$\begin{aligned}\left\langle \mathbf{G}^A \left| \mathbf{G}^B \right\rangle &= \int d\mathbf{r}_1 \mathbf{G}^A(\mathbf{r}_1) r_{12}^{-1} \mathbf{G}^B(\mathbf{r}_1) \\ &\cong \frac{\mu_T^A \mu_T^B}{R_{AB}^3} (\sin\theta_A \sin\theta_B - 2\cos\theta_A \cos\theta_B)\end{aligned} \tag{11}$$

where μ_T^A is the transition dipole. Here the angle θ_A is between the vectors $\mathbf{R}_{AB}$ and μ_T^A so that it satisfies $\mu_T^A R_{AB} \cos\theta_A = \mu_T^A \cdot \mathbf{R}_{AB}$. This interaction is largest in magnitude when the transition dipoles are parallel or anti–parallel along $\mathbf{R}_{AB}$, and there are many configurations where the interaction is also significant.

Determining the relative contribution to a particular singular value of the σ versus π–bonds is also possible with the second unfolding of the $\boldsymbol{T}_2^{disp}$ amplitudes. Four general cases are present when considering the relative contributions of the σ and π–bonds: both the π and σ related virtual orbitals contribute approximately equally to dispersion; only the σ related virtual orbitals contribute significantly; only the π related virtual orbitals contribute significantly; and both the σ and π related virtual orbitals do not contribute significantly.

These cases can be understood with visualization of the virtual orbitals. For the first case, in which both the π and σ–bonds contribute significantly to dispersion, the virtual orbitals associated with the σ–bonds (p–like) and π–bonds (d–like) align well parallelly between the two distant bonds, as shown in figure **??**. This leads to an enhanced dispersion contribution. For the second case, in which the σ–bond dominates the dispersion contribution, the virtual orbitals associated with the σ–bonds align well, while the virtual orbitals associated with the π–bonds align perpendicularly, as shown in figure 14. In the third case, in which the π–bond dom-

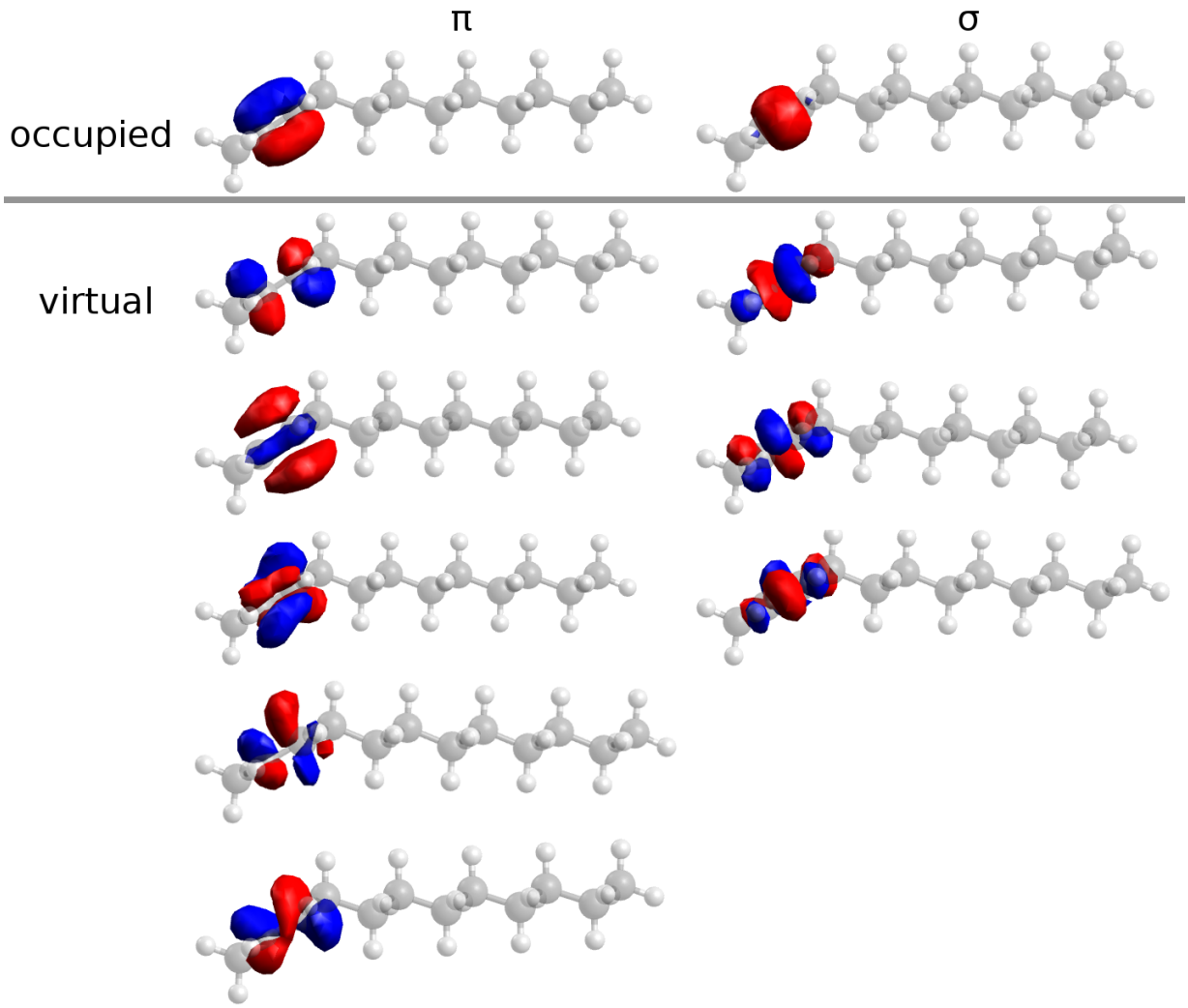

FIG. 12: Occupied π and σ–orbitals and corresponding virtual components of the first five/three important geminals for dispersion of trans–2–dodecene. Note that the most important virtual orbital components are simply one angular momentum higher than the corresponding occupied orbitals, i.e., one node is introduced.

inates the dispersion contribution, the virtual orbitals associated with the π–bonds align parallelly, while the virtual orbitals associated with the σ–bonds align perpendicularly, as shown in figure 15. In this case, one would expect the π related (d–like) virtual dispersion interactions to be insignificant, since they would contribute only a quadrupole–quadrupole term. However, it appears they can still "carry–the–weight" of dispersion even when the virtual orbitals associated with the σ–bonds (p–like) align poorly.

3. Redundancy of virtual orbitals

Examination of the second decomposition of the geminals into occupied–virtual sets also allows for the exploration of redundancy in the dispersion–specific virtual space. To test for redundancies, the set of the three most important virtual orbitals on CC–bond 1 were used to construct an overlap matrix. This set of fifteen virtual orbitals centered on CC–bond 1 represent the most important virtual orbitals for interactions between CC–bond 1 and CC–bonds 3, 5, 7, and 9. The eigenvalues of the overlap matrix are then used to determine the linear dependence of the set. In the worst case scenario all virtual orbitals would be linearly independent, resulting in fifteen eigenvalues equaling exactly one. In the best case scenario only three virtual orbitals would be linearly independent; this would result in three eigenvalues equaling precisely 5, with the remaining equalling 0. Preforming such analysis on n–dodecane results in the eigenvalues presented in table II. As can be seen, the best case scenario of linear dependence is approached, but not reached. However, exact reproduction of linear dependence is not expected since these dispersion–specific virtuals were not constructed with this feature in mind. On

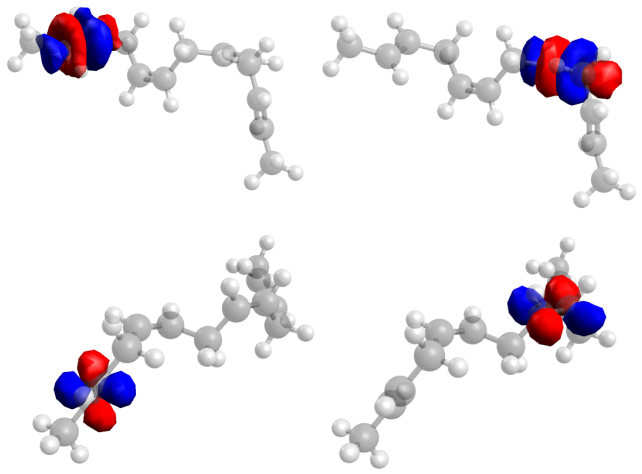

FIG. 13: The virtual components of the first geminal of the first and third CC–double–bond of 2,5,8,11–tridecene. The top panel shows the virtual orbitals associated with the σ–bond dispersion, and the bottom panel shows the virtual orbitals associated with the π–bond dispersion. Good alignment is observed for both σ and π–bond associated virtual orbitals.

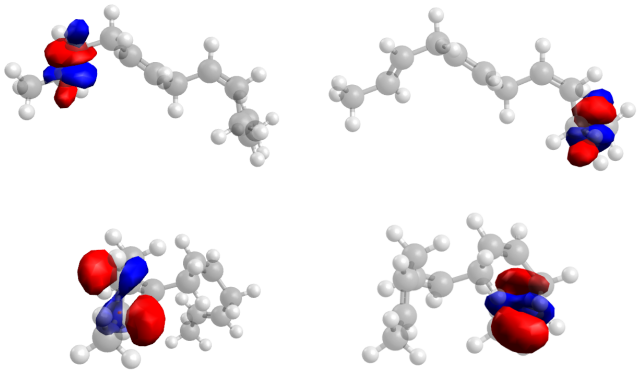

FIG. 14: The virtual components of the second geminal of the first and fourth CC–double–bond of 2,5,8,11–tridecene. The top panel shows the virtual orbitals associated with the σ–bond dispersion, and the bottom panel shows the virtual orbitals associated with the π–bond dispersion. Good alignment is observed for the σ–bond associated virtual orbitals, while poor alignment is observed for the π associated virtual orbitals.

this basis however, the level of linear dependence they do reach is promising for the possibility of constructing a set of $3N$ dispersion–specific virtual orbitals, rather than $3N(N-1)$.

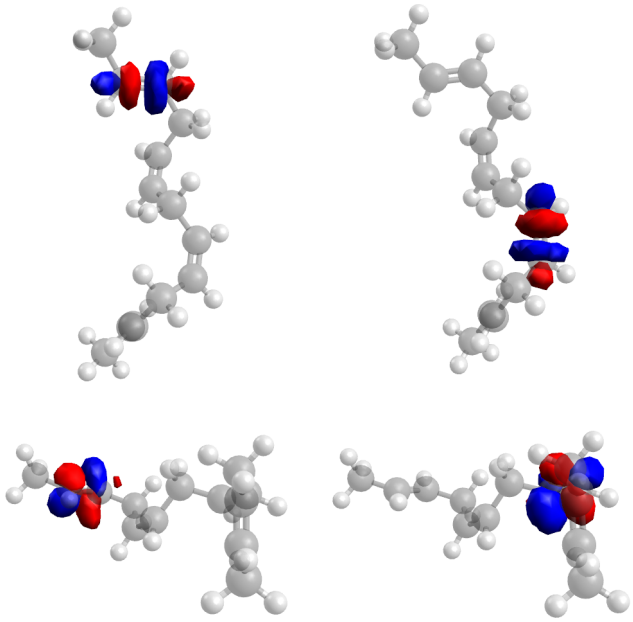

FIG. 15: The virtual components of the third geminal of the first and third CC–double–bond of 2,5,8,11–tridecene. The top panel shows the virtual orbitals associated with the σ–bond dispersion, and the bottom panel shows the virtual orbitals associated with the π–bond dispersion. Good alignment is observed for the π–bond associated virtual orbitals, while poor alignment is observed for the σ–bond associated virtual orbitals.

V. CONCLUSIONS

In this work the feasibility of compressing the CCSD amplitudes associated with dispersion ($\boldsymbol{T}_2^{disp}$) was explored. SVD analysis of the $\boldsymbol{T}_2^{disp}$ amplitudes in a many–body (He trimer) pairwise dispersion case shows the number of important geminals stays fixed at three per pairwise interacting body. That is to say no additional geminals of equal magnitude to the dimer interactions are introduced. This allows for a simplified pairwise SVD analysis of interacting fragments, with the influence of the remaining fragments entering as three–body dispersion through simply the relaxation of the original $\boldsymbol{T}_2$ amplitudes. This premise allows for an uncomplicated SVD analysis of the $\boldsymbol{T}_2^{disp}$ amplitudes in an intramolecular setting, where the number of dispersion interactions escalate quickly with molecular size.

Five long–chain hydrocarbons and one silane molecule were chosen for the intramolecular SVD analysis. Consistent with the previous work on dimers[25], three geminals dominate the contribution to dispersion in the hydrocarbons at long distances. For single bonded systems, at a bond–bond distance of approximately 5 Å the drop–off between the third and fourth singular values approximately one order of magnitude in difference, with the

drop–off between the first and fourth singular value being nearly two orders of magnitude in difference. At distances greater than 5.1 Å the compressing of the virtual space to three orbitals per occupied captures approximately 99% of the dispersion energy, with the error dropping rapidly beyond such distances. Conversely, at distances between 2.6 and 5.1 Å three virtual orbitals capture approximately 96% of the dispersion energy. However, these errors are expected at these small distances due to the neglect of short–range exchange effects. For double–bonded systems a similar drop–off is observed between the third and *fifth* singular value, however the decay with distance of the fourth singular value is slower than is observed in the single bonded cases. This slower decay was found to be due to a dipole–quadrupole interactions, and suggests the necessity of including at least a fourth dispersion specific virtual orbital for each occupied π–orbital.

Decay rates for the first three singular values of the hydrocarbons were found to go as approximately R^{-3} (R^{-6} in energy), consistent with the dipole–dipole terms of a dispersion interaction. However, the first singular value tends to decay closer to R^{-2}, possibly due to through–bond enhancement of dispersion. It was found that this enhancement lowers the decay rate for the total intramolecular dispersion energies from the expected R^{-6} to $R^{-5.7}$. The remaining singular values considered (four through seven) appear to decay approximately as R^{-5} (R^{-10} in energy), consistent with the quadrupole–quadrupole terms of dispersion interactions. Decay rates closer to R^{-6} were occasionally observed, suggesting higher multi–pole interactions. However, this could not be confirmed and may be simply due to statistical errors, or even relative orientation of the interacting virtual orbitals. Interestingly, dipole–quadrupole R^{-4} (R^{-8} in energy) decay rates were only observed to be significant contributors to dispersion in the double bonded systems. These being represented by the fourth singular values in the analysis. These dipole–quadrupole interactions are found to occur between the virtual orbitals associated with the π–electrons and the virtual orbitals associated with the σ–electrons.

A Boys localization of the occupied and virtual space provided a sensible choice of bond–specific orbitals for use in the SVD analysis. However, what constituted a fragment within a given molecule was open for interpretation. Two schemes were chosen to test the significance of this choice: bond–bond dispersion interactions (CC bonds) and chemical group dispersion interactions (CH_x groups). n–Hexane and Si_{10}-silane both were analyzed with each of these schemes. Both the bond–bond and group dispersion interactions are consistent with the conclusion that a significant drop off in the singular values is observed between the third and fourth singular value. This suggests that it may be possible to further compress the necessary dispersion specific virtual orbitals to be functional group specific rather than bond specific. The impact this would have on the resulting dispersion energies is currently being explored.

Analysis of the virtual orbitals associated with the highest singular values show an interesting trend; they resemble orbitals of one angular momentum higher than the corresponding occupied orbital, i.e., the first three largest geminals coincide with an occupied σ–orbital corresponding to a set of three p–like virtual orbitals, and likewise the first five largest geminals coincide with an occupied π–orbital corresponding to a set of five d–like virtual orbitals. The simplicity of this correspondence may aid in the construction of dispersion–specific virtual orbitals. Decay rates suggest an inclusion of at least three dispersion–specific virtual orbitals per occupied σ–orbital, consistent with the analysis of the virtual orbitals. Decay rates also suggest the need for at least four dispersion–specific virtual orbitals per occupied π–orbital, however analysis of the virtual orbitals suggests five virtual d–like orbitals may be necessary depending on relative orientation of the bonds within a molecule.

Much discussion in the literature has been given to the relative importance of σ versus π–electrons in dispersion interactions. In this work orientation was found to play a strong role when determining which of the two is the larger contributor. In cases where the orientation of the two interacting bonds were favorable to aligning both the p–like and d–like virtual orbitals, both the σ and π–electrons contributed similar amounts. However, when the orientation of the two interacting bonds were such that the two p–like virtual orbitals align well, but the d–like virtual orbitals did not, the σ–electrons were found to dominate dispersion between the two bonds. Conversely, when d–like virtual orbitals aligned well and the p–like virtual orbitals did not, the π–electrons were found to dominate dispersion.

Lastly, the dispersion–specific virtual space were shown to, at least to a significant part, contain some amount of linear dependence. This may allow for the construction of a small set of versatile dispersion–specific orbitals that can account for *all* dispersion interactions rather than the need for a larger set to account for each *individual* dispersion interaction (i.e., $3N$ rather than $3N(N-1)$ virtuals).

The compressibility of the dispersion specific T_2^{disp} amplitudes will mark a dramatic speed–up in methods such as rung–5 DFT calculations based on RPA, especially when coupled with other methods such as tensor hyper–contraction (THC) or interpolative separable density fitting decomposition (IDSF). Compared with a full CCSD calculation (as far as dispersion interactions are concerned) the speed–up could be as high as

$$\left(\frac{N_{Ao} - N_{occ}}{3N_{occ}}\right)^4$$

where N_{Ao} is the size of the atomic orbital basis, and N_{occ} is the number of active occupied orbitals.

ACKNOWLEDGMENTS

This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Advanced Scientific Computing Research and Office of Basic Energy Sciences, Scientific Discovery through Advanced Computing (SciDAC) program.

[1] S. Kristyn and P. Pulay, Chem. Phys. Lett. **229**, 175 (1994).
[2] M. Dion, H. Rydberg, E. Schröder, D. C. Langreth, and B. I. Lundqvist, Phys. Rev. Lett. **92**, 246401 (2004).
[3] O. A. Vydrov and T. Van Voorhis, J. Chem. Phys. **133**, 244103 (2010).
[4] S. Grimme, J. Antony, S. Ehrlich, and H. Krieg, J. Chem. Phys. **132**, 154104 (2010).
[5] S. Grimme, S. Ehrlich, and L. Goerigk, J. Comput. Chem. **32**, 1456 (2011).
[6] A. Tkatchenko and M. Scheffler, Phys. Rev. Lett. **102**, 073005 (2009).
[7] A. D. Becke and E. R. Johnson, J. Chem. Phys. **123**, 154101 (2005).
[8] A. D. Becke and E. R. Johnson, J. Chem. Phys. **124**, 014104 (2006).
[9] M. Del Ben, O. Schütt, T. Wentz, P. Messmer, J. Hutter, and J. Van de Vondele, Comput. Phys. Commun. **187**, 120 (2015).
[10] J. Cizek, Adv. Chem. Phys **14**, 35 (1966).
[11] C. Ochsenfeld, J. Kussmann, and D. S. Lambrecht, Reviews in Computational Chemistry **23**, 1 (2007).
[12] B. W. Hopkins and G. S. Tschumper, J. Phys. Chem. A **108**, 2941 (2004).
[13] X. Ren, P. Rinke, C. Joas, and M. Scheffler, J. Mater. Sci. **47**, 7447 (2012).
[14] S. Kurth and J. P. Perdew, Phys. Rev. B **59**, 10461 (1999).
[15] Z. Yan, J. P. Perdew, and S. Kurth, Phys. Rev. B **61**, 16430 (2000).
[16] J. P. Perdew and S. Kurth, "Density functionals for non-relativistic coulomb systems in the new century," in *A Primer in Density Functional Theory*, edited by C. Fiolhais, F. Nogueira, and M. A. L. Marques (Springer Berlin Heidelberg, Berlin, Heidelberg, 2003) pp. 1–55.
[17] J. P. Perdew, A. Ruzsinszky, J. Tao, V. N. Staroverov, G. E. Scuseria, and G. I. Csonka, J. Chem. Phys. **123**, 062201 (2005).
[18] N. Mardirossian and M. Head-Gordon, J. Chem. Phys. **148**, 241736 (2018).
[19] G. Santra, N. Sylvetsky, and J. M. L. Martin, J. Phys. Chem. A **123**, 5129 (2019).
[20] C. Riplinger and F. Neese, J. Chem. Phys. **138**, 034106 (2013).
[21] M. Schütz and H.-J. Werner, J. Chem. Phys. **114**, 661 (2001).
[22] S. Saebo and P. Pulay, Annual Review of Physical Chemistry **44**, 213 (1993).
[23] S. Ten-no, Chem. Phys. Lett. **398**, 56 (2004).
[24] D. P. Tew and W. Klopper, J. Chem. Phys. **123**, 074101 (2005).
[25] J. F. Gonthier and M. Head-Gordon, J. Chem. Phys. **147**, 144110 (2017).
[26] P. Pulay, Chem. Phys. Lett. **100**, 151 (1983).
[27] A. J. Stone, *The Theory of Intermolecular Forces* (Oxford University Press, Oxford, 1996).
[28] P. Jurecka, J. Sponer, J. Cerny, and P. Hobza, Phys. Chem. Chem. Phys. **8**, 1985 (2006).
[29] S. M. Cybulski and M. L. Lytle, J. Chem. Phys. **127**, 141102 (2007).
[30] M. Schütz, G. Rauhut, and H.-J. Werner, J. Phys. Chem. A **102**, 5997 (1998).
[31] R. J. Azar and M. Head-Gordon, J. Chem. Phys. **136**, 024103 (2012).
[32] W. B. Schneider, G. Bistoni, M. Sparta, M. Saitow, C. Riplinger, A. A. Auer, and F. Neese, J. Chem. Theory Comput. **12**, 4778 (2016).
[33] T. Kinoshita, O. Hino, and R. J. Bartlett, The Journal of Chemical Physics **119**, 7756 (2003).
[34] R. Krishnan, M. J. Frisch, and J. A. Pople, J. Chem. Phys. **72**, 4244 (1980).
[35] G. Chalasinski and M. M. Szczesniak, Chem. Rev. **94**, 1723 (1994).
[36] G. Chałasiński and M. M. Szcześniak, Chem. Rev. **100**, 4227 (2000).
[37] M. J. Elrod and R. J. Saykally, Chem. Rev. **94**, 1975 (1994).
[38] J. F. Gonthier and M. Head-Gordon, J. Chem. Theory Comput. **15**, 4351 (2019).
[39] K. Szalewicz, R. Bukowski, and B. Jeziorski, in *Theory and Applications of Computational Chemistry*, edited by C. E. Dykstra, G. Frenking, K. S. Kim, and G. E. Scuseria (Elsevier, Amsterdam, 2005) pp. 919 – 962.
[40] S. Wen and G. J. O. Beran, J. Chem. Theory Comput. **8**, 2698 (2012).
[41] S. Wen and G. J. O. Beran, J. Chem. Theory Comput. **7**, 3733 (2011).
[42] R. Podeszwa, B. M. Rice, and K. Szalewicz, Phys. Rev. Lett. **101**, 115503 (2008).
[43] B. Temelso, C. R. Renner, and G. C. Shields, J. Chem. Theory Comput. **11**, 1439 (2015).
[44] R. L. Martin, J. Chem. Phys. **118**, 4775 (2003).
[45] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).
[46] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).
[47] N. Mardirossian and M. Head-Gordon, Phys. Chem. Chem. Phys. **16**, 9904 (2014).
[48] P. C. Hariharan and J. A. Pople, Theor. Chim. Acta **28**, 213 (1973).
[49] G. D. Purvis and R. J. Bartlett, J. Chem. Phys. **76**, 1910 (1982).
[50] T. H. Dunning Jr, J. Chem. Phys. **90**, 1007 (1989).
[51] Y. Shao, Z. Gan, E. Epifanovsky, A. T. Gilbert, M. Wormit, J. Kussmann, A. W. Lange, A. Behn, J. Deng, X. Feng, D. Ghosh, M. Goldey, P. R. Horn, L. D. Jacobson, I. Kaliman, R. Z. Khaliullin, T. Kus, A. Landau, J. Liu, E. I. Proynov, Y. M. Rhee, R. M. Richard, M. A. Rohrdanz, R. P. Steele, E. J. Sundstrom, H. L. Woodcock, III, P. M. Zimmerman, D. Zuev, B. Albrecht, E. Alguire, B. Austin, G. J. O. Beran, Y. A. Bernard, E. Berquist, K. Brandhorst, K. B. Bravaya, S. T. Brown, D. Casanova, C.-M. Chang, Y. Chen, S. H. Chien, K. D. Closser, D. L. Crittenden, M. Diedenhofen, R. A. DiStasio, Jr., H. Do, A. D. Dutoi, R. G. Edgar, S. Fatehi, L. Fusti-Molnar, A. Ghysels, A. Golubeva-Zadorozhnaya, J. Gomes, M. W. Hanson-Heine, P. H. Harbach, A. W. Hauser, E. G. Hohenstein, Z. C. Holden, T.-C. Jagau, H. Ji, B. Kaduk, K. Khistyaev, J. Kim, J. Kim, R. A. King, P. Klunzinger, D. Kosenkov, T. Kowalczyk, C. M. Krauter, K. U. Lao, A. D. Laurent, K. V. Lawler, S. V. Levchenko, C. Y. Lin, F. Liu, E. Livshits, R. C. Lochan, A. Luenser, P. Manohar, S. F. Manzer, S.-P. Mao, N. Mardirossian, A. V. Marenich, S. A. Maurer, N. J. Mayhall, E. Neuscamman, C. M. Oana, R. Olivares-Amaya, D. P. O'Neill, J. A. Parkhill, T. M. Perrine, R. Peverati, A. Prociuk, D. R. Rehn, E. Rosta, N. J. Russ, S. M. Sharada, S. Sharma, D. W. Small, A. Sodt, T. Stein, D. Stuck, Y.-C. Su, A. J. Thom, T. Tsuchimochi, V. Vanovschi, L. Vogt, O. Vydrov, T. Wang, M. A. Watson, J. Wenzel, A. White, C. F. Williams, J. Yang, S. Yeganeh, S. R. Yost, Z.-Q. You, I. Y. Zhang, X. Zhang, Y. Zhao, B. R. Brooks, G. K. Chan, D. M. Chipman, C. J. Cramer, W. A. Goddard, III, M. S. Gordon, W. J. Hehre, A. Klamt, H. F. Schaefer, III, M. W. Schmidt, C. D. Sherrill, D. G. Truhlar, A. Warshel, X. Xu, A. Aspuru-Guzik, R. Baer, A. T. Bell, N. A. Besley, J.-D. Chai, A. Dreuw, B. D. Dunietz, T. R. Furlani, S. R. Gwaltney, C.-P. Hsu, Y. Jung, J. Kong, D. S. Lambrecht, W. Liang, C. Ochsenfeld, V. A. Rassolov, L. V. Slipchenko, J. E. Subotnik, T. V. Voorhis, J. M. Herbert, A. I. Krylov, P. M. Gill, and M. Head-Gordon, Mol. Phys. **113**, 184 (2015).
[52] S. F. Boys, Rev. Mod. Phys. **32**, 296 (1960).
[53] S. F. Boys, *Quantum Science of Atoms, Molecules, and Solids*, Academic Press, NY , 253 (1966).

[54]J. E. Subotnik, Y. Shao, W. Liang, and M. Head-Gordon, J. Chem. Phys. **121**, 9220 (2004).

[55]J. E. Subotnik, A. Sodt, and M. Head-Gordon, Phys. Chem. Chem. Phys. **9**, 5522 (2007).